

---

# **Evaluación estadística de las observaciones de ozono suministradas por la estación ambiental Catumare**

---



Presentado por

**Martín Andrés Suárez Pintor**

**LOS LIBERTADORES**  
FUNDACIÓN UNIVERSITARIA

**Fundación Universitaria Los Libertadores**

Facultad de Ingeniería y Ciencias Básicas

Especialización en Estadística Aplicada

Bogotá D.C, Colombia

2020

---

**Evaluación estadística de las  
observaciones de ozono  
suministradas por la estación  
ambiental Catumare**

---

**LOS LIBERTADORES**

Presentado por  
**Martín Andrés Suárez Pintor**

en cumplimiento parcial de los requerimientos para optar al título

de

**Especialista en Estadística Aplicada**

*Quien asesora*

**José John Fredy González Veloza**

**Fundación Universitaria Los Libertadores**

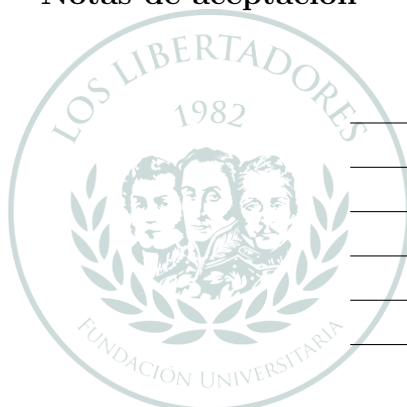
Facultad de Ingeniería y Ciencias Básicas

Especialización en Estadística Aplicada

Bogotá D.C, Colombia

2020

## Notas de aceptación



# LOS LIBERTADORES

## FUNDACIÓN UNIVERSITARIA

---

Firma del presidente del jurado

---

Firma del jurado

---

Firma del jurado

Bogotá D.C., Abril de 2020.



# LOS LIBERTADORES

## FUNDACIÓN UNIVERSITARIA

Las directivas de la Fundación Universitaria Los Libertadores, los jurados calificadores y el cuerpo docente no son responsables por los criterios e ideas expuestas en el presente documento. Estos corresponden únicamente a los autores y a los resultados de su trabajo.

# Índice general

<b>1</b>	<b>Introducción</b>	<b>9</b>
<b>2</b>	<b>Planteamiento del Problema</b>	<b>11</b>
	Objetivos . . . . .	12
	Objetivo General . . . . .	12
	Objetivos Específicos . . . . .	12
	Justificación . . . . .	12
<b>3</b>	<b>Marco Teórico</b>	<b>15</b>
	Contaminación atmosférica . . . . .	16
	Contaminación en el Aire . . . . .	16
	Tipos de Contaminantes . . . . .	16
	Material Particulado . . . . .	17
	Gases . . . . .	17
	Ozono Troposférico . . . . .	18
	Normatividad Colombiana Aplicable . . . . .	19
	Índice de Calidad del Aire – ICA . . . . .	20
	Estadística aplicada . . . . .	21
	Muestreo Probabilístico . . . . .	22
	Población . . . . .	22
	Muestra . . . . .	22
	Muestreo . . . . .	22
	Tamaño de la Muestra . . . . .	22
	Tamaño de la Muestra según la proporción . . . . .	23
	Tipo de Muestreo . . . . .	23
	Muestreo Probabilístico Sistemático . . . . .	23
	Análisis Multivariado . . . . .	24
	Análisis de Componentes Principales . . . . .	24
	Control Estadístico de la Calidad . . . . .	24
	Gráficos de control . . . . .	25

<b>4</b>	<b>Diseño Metodológico</b>	<b>27</b>
	Nombre del capítulo . . . . .	27
	Método . . . . .	27
	Enfoque . . . . .	27
	Diseño Metodológico . . . . .	27
	Fase 1. Fuente y organización de la información . . . . .	27
	Base de datos . . . . .	28
	Datos Empleados para el estudios . . . . .	28
	Fase 2. Análisis Estadístico . . . . .	28
	Organización de la base de datos . . . . .	28
	Análisis Descriptivo . . . . .	28
	Análisis Multivariado . . . . .	29
	Cartas de Control Estadístico . . . . .	29
	Fase 3. Análisis de Resultados . . . . .	29
<b>5</b>	<b>Análisis y Resultados</b>	<b>31</b>
	Organización y Toma de Muestra . . . . .	31
	Análisis Descriptivo Univariado . . . . .	32
	Medidas de tendencia central, dispersión y forma . . . . .	33
	Test de distribución normal univariado . . . . .	34
	Análisis Descriptivo Multivariado . . . . .	35
	Análisis de Correlación . . . . .	35
	Test de distribución Normal Multivariado . . . . .	36
	Distancia generalizada . . . . .	37
	Test de Multinomialidad: Prueba de Mardia . . . . .	37
	Análisis de Componentes Principales (ACP) . . . . .	38
	Cartas de Control Estadístico . . . . .	39
	Cartas de Control univariada . . . . .	40
	Cartas de Control Multivariado . . . . .	42
	Carta de Control en Base a Análisis de Componentes Principales . . . . .	42
<b>6</b>	<b>Conclusiones</b>	<b>45</b>
	<b>Referencias</b>	<b>46</b>

# Índice de figuras

<b>5-1</b>	Selección de la cantidad de muestra a utilizar por semana. La mínima cantidad de observaciones por muestra se da en el día seis con un total de 48. *Margen de error y **Nivel de confianza para una proporción de éxito ( $p = 0.5$ ) y de fracaso ( $q = 0.5$ ) como se observa en la ecuación 3-3 . . . . .	32
<b>5-2</b>	Análisis univariado para el contaminante atmosférico Ozono . . . . .	35
<b>5-3</b>	Matriz de correlación de Pearson . . . . .	36
<b>5-4</b>	Determinación gráfica de distribución normal multivariable por medio de la distribución chi-cuadrado con seis grados de libertad . . . . .	37
<b>5-5</b>	Análisis multivariado para los dos primeros componentes principales . . . . .	39
<b>5-6</b>	Carta de control estadístico CUSUM para el contaminante atmosférico Ozono	40
<b>5-7</b>	Carta de control estadístico EWMA para el contaminante atmosférico Ozono	41
<b>5-8</b>	Control de calidad multivariado para los dos primeros componentes principales . . . . .	42

# Índice de tablas

<b>3-1</b>	Contaminantes atmosféricos particulados. Fuente: <a href="#">IDEAM (2018)</a> . . . . .	17
<b>3-2</b>	Contaminantes atmosféricos gaseosos. Fuente: <a href="#">IDEAM (2018)</a> . . . . .	18
<b>3-3</b>	Normatividad de calidad del aire en Colombia vrs. Recomendaciones de la Organización Mundial de la Salud. Las unidades correspondientes a cada uno de los contaminantes son, para $PM_{10} \mu g/m^3$ , $PM_{2.5} \mu g/m^3$ , $NO_2$ ppm, $SO_2$ ppm, $O_3$ ppm, CO ppm Fuente: <a href="#">IDEAM (2018)</a> . . . . .	19
<b>3-4</b>	Puntos de corte del Índice de Calidad del Aire ICA en Colombia para material particulado y dióxido de carbono. Fuente: <a href="#">IDEAM (2018)</a> . . . . .	20
<b>3-5</b>	Puntos de corte del Índice de Calidad del Aire ICA en Colombia para gases atmosféricos contaminantes. Fuente: <a href="#">IDEAM (2018)</a> . . . . .	21
<b>5-1</b>	Análisis descriptivo univariado del promedio del muestreo de 21 semanas *	
	Rango Inter Cuartílico. . . . .	33
<b>5-2</b>	Resultado por variable del test de normalidad de Shapiro Wilk . . . . .	34
<b>5-3</b>	Prueba de Mardia para determinar multinormalidad . . . . .	37
<b>5-4</b>	Contribución de las variables a las dimensiones o componentes principales .	38





# Evaluación estadística de las observaciones de ozono suministradas por la estación ambiental Catumare

## Resumen

El ozono es considerado el tercer contaminante atmosférico de estudio detrás de los materiales particulados. Con el fin de seguir la recomendación de la Organización Internacional de la Salud, el gobierno nacional por medio del Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM) realiza el control de contaminantes atmosféricos a través de unidades ambientales ubicadas en zonas urbanas y de minería. En Villavicencio, el IDEAM entregó a cargo de Cormacarena la estación ambiental Catumare, la cual realiza mediciones de Ozono ( $O_3$ ) y variables meteorológicas como velocidad del viento, dirección del viento, temperatura, humedad, radiación solar, presión atmosférica y precipitación. Debido a que el 36 % de los datos fueron cero, se hizo necesario realizar una depuración de la base de datos, logrando obtener una muestra de 48 datos de cada una de las variables por cada una de las 21 semanas, para aplicar cada uno de los análisis estadísticos. Entre los análisis se encuentra análisis descriptivo univariado y multivariado, arrojando correlaciones positivas entre el ozono y la temperatura, velocidad del viento y la radiación; mientras que la presión y la humedad presentaron correlaciones negativas. Las cartas de control univariado para el ozono evidenció semanas con observaciones fuera de control. No obstante, la carta de multivariado dejó en evidencia que se trataba de variables aleatorias o naturales las posibles causas de alteración en los datos del ozono.

**Palabras claves:** Ozono troposférico, Calidad del aire, Análisis de componentes principales, Cartas de control estadístico, variables meteorológicas.



# Capítulo 1

## Introducción

Los contaminantes atmosféricos cada día toman una mayor relevancia. Años atrás se hablaba de la contaminación en los recursos hídricos y la contaminación que se enfrenta por los residuos sólidos. El gobierno nacional por medio del Ministerio del Medio Ambiente, han acatado las recomendaciones de la Organización Mundial de la Salud. Según la Organización Mundial de la salud (OMS), el aire limpio es una necesidad para el bienestar y la salud de la humanidad. Sin embargo, este mismo organismo internacional estima que una de cada nueve muertes en el mundo está relacionada con la contaminación atmosférica ([Organization y cols., 2016](#)). Entre los contaminantes más relevantes, encontramos los obtenidos del proceso de combustión de combustibles fósiles. La OMS hace énfasis en material particulado (PM) y dos límites máximos permitidos en los que basan su estándar de prevención. Para  $PM_{10}$  (partículas de  $10\mu m$  o menos) es  $20\mu g/m^3$  y una exposición máxima de  $10\mu g/m^3$  para las  $PM_{2.5}$  (partículas de  $2.5\mu m$  o menos) ([Organization y cols., 2006](#)). Por medio del Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM), se realiza el registro y control del medio ambiente en zonas urbanas del territorio nacional. Entre las más controladas tenemos Bogotá, Medellín, Yumbo, Santa Marta, Cali, Nemocón, Ráquira, Sibaté y en la Zona Minera del Cesar, pero de manera constante busca llegar a nuevos lugares del país para realizar un cubrimiento mayor de la población colombiana. En el departamento del Meta se encuentra el municipio de Villavicencio donde el IDEAM tiene instalada tres unidades móviles de control ambiental a cargo de la Corporación Para El Desarrollo Sostenible del Área Manejo Especial la Macarena (Cormacarena), ubicados en Cofren, La Esmeralda y Catumare, allí se registra la concentración de diversos contaminantes atmosféricos, dentro de ellos el ozono. El ozono es la tercera sustancia de mayor estudio en el país. Se le llama también como Ozono troposférico debido a que es originado en la capa más baja de la atmosfera. Su capacidad de disolución permite la penetración en las vías respiratorias desembocando en la irritación de los tejidos pulmonares y mucosas. Su formación está dada por medio de una reacción fotoquímica en la que interviene el metano, monóxido de nitrógeno y monóxido de

carbono, los cuales forman radicales de oxígeno que después reaccionarán con el oxígeno atmosférico y formará el ozono. Las diferentes características, condiciones de formación, tipos de contaminantes y normatividad asociada podrá ser encontrada en el capítulo 3. El gobierno colombiano por medio de la página web Datos Abiertos, administrada por el Ministerio de Tecnologías de la Información y las Comunicaciones, deja a disposición del interesado bases de datos de entidades públicas con múltiples temas de interés nacional, regional o local. Para este trabajo se empleó la base de datos que lleva por nombre “Datos de calidad del aire en la estación Catumare Junio a Diciembre 2018”. El enlace y demás información de la base de datos, los encontrará en el capítulo 4. Por lo tanto, se plantea la siguiente hipótesis de este trabajo: **Las inconsistencias de los datos obtenidos en la estación ambiental Catumare provoca incorrectos análisis estadístico del contaminante atmosférico ozono.** Para realizar un análisis estadístico de la base de datos es necesario realizar un análisis descriptivo univariado de la base de datos; como también, un análisis descriptivo multivariado con sus respectivas pruebas de hipótesis de distribución con el fin de aplicar un método paramétrico o no paramétrico. Además, de realizar una evaluación del comportamiento de los datos por medio de gráficas de control de calidad univariado y multivariado. El capítulo 5 mostrara los diferentes desempeños y resultados de las variables de estudio.

## Capítulo 2

### Planteamiento del Problema

El gobierno nacional, por medio del Ministerio de Ambiente y Desarrollo realiza el seguimiento correspondiente a la calidad del aire buscando dar cumplimiento al artículo 79 y 80 de la Constitución Política de Colombia (1991). Los cuales consagran el derecho colectivo a gozar de un ambiente sano y, como deber del estado, prevenir los factores de deterioro ambiental. En el literal a), artículo 8 del Decreto Ley 2811 de 1974, hace referencia que la contaminación en el aire es uno de los factores que deterioran el ambiente. Además, que los artículos 2.2.5.1.2.4, 2.2.5.1.2.8 y los literales b), j) y k) del artículo 2.2.5.1.6.1 del Decreto 1076 de 2015, fija la competencia del Ministerio de Ambiente y Desarrollo Sostenible para establecer las normas de calidad de aire o inmisión, los niveles periódicos de inmisión, la prevención y control de la contaminación atmosférica proveniente de actividades mineras, industriales y de transporte. Esta última, de gran impacto en la actualidad atmosférica de las ciudades a nivel nacional. Según la Organización Mundial de la salud (OMS), el aire limpio es una necesidad para el bienestar y la salud de la humanidad. Sin embargo, este mismo organismo internacional estima que una de cada nueve muertes en el mundo está relacionada con la contaminación atmosférica (Organization y cols., 2016). Más de dos millones de muertes prematuras lo atribuyen a la contaminación atmosférica por año y más del 50 % de las muertes se dan en los países en vía de desarrollo. (Organization, 2006).

Entre los contaminantes más relevantes, encontramos los obtenidos del proceso de combustión de combustibles fósiles. La OMS hace énfasis en material particulado (PM) y dos límites máximos permitidos en los que basan su estándar de prevención. Para  $PM_{10}$  (partículas de  $10\mu m$  o menos) es  $20\mu g/m^3$  y una exposición máxima de  $10\mu g/m^3$  para las  $PM_{2.5}$  (partículas de  $2.5\mu m$  o menos)(Organization y cols., 2006).

Esta relevante necesidad de mantener un constante monitoreo de la contaminación atmosférica, da pie a que el Ministerio de Ambiente por medio del Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM) realicen toma de datos a nivel nacional. Estos datos son analizados para presentar un informe anual de los contaminantes contro-

lados según la resolución 2254 del 2017([de Ambiente y Desarrollo Sostenible, 2017](#)).

En el 2018 el IDEAM presentó el informe correspondiente a la vigencia 2017. En este, se dice que el 88 % de las estaciones de monitoreo reportaron cumplimiento de la norma, en comparación al 2016 donde solo se alcanzó el 65.6 %. Siendo las estaciones de Bogotá y Medellín las que presentaron los niveles más altos ([IDEAM, 2018](#)). Por lo tanto se realiza la siguiente pregunta de investigación: **¿La base de datos suministrada por la estación ambiental Catumare entrega la suficiente información para realizar el análisis estadístico del contaminante atmosférico ozono?**

## Objetivos

### Objetivo General

Evaluar las observaciones del contaminante ambiental ozono suministrados por la estación ambiental de Catumare del IDEAM en el municipio de Villavicencio por medio de análisis de componentes principales (ACP) y control estadístico de la calidad.

### Objetivos Específicos

- Depurar la base de datos excluyendo los valores iguales a cero de acuerdo a lo que establece el criterio del IDEAM y dejándola organizada por el promedio semanal de cada variable.
- Aplicar análisis de componentes principales a los datos seleccionados de ozono y condiciones meteorológicas con el fin de encontrar sus posibles relaciones.
- Realizar cartas de control a las observaciones de ozono para detectar posibles semanas fuera de control y su relación con las variables meteorológicas.

## Justificación

Para las entidades gubernamentales la información es relevante para la toma de decisiones, una entidad sin la información correcta, oportuna y de calidad, carecerá de las herramientas que la lleven a una toma de acciones frente a los múltiples desafíos y deberes a la que es sometida a diario. Ese deber que tienen las entidades a nivel nacional e internacional por garantizar sistemas de monitoreo de los contaminantes ambientales y la información del comportamiento que estos tienen con el fin de tomar medidas correctas frente a la eliminación o, como mínimo, la disminución de este material particulado o gaseoso, refleja la necesidad de plantear un confiable análisis estadístico que permita la correcta recolección y su posterior escudriñamiento.

Para realizar un correcto análisis estadístico, se debe partir de una base de datos confiable que permita interpretar los diferentes factores asociados, posibles comportamientos o

condición actual y futura de las variables estudiadas. Por lo tanto, se debe garantizar la correcta toma de datos. Por tal motivo es de vital importancia que el primer paso para el análisis estadístico, después de la recolección y tabulación de las observaciones, sea la evaluación de datos al comprobar su comportamiento normalizado. Y dado el caso, indagar, analizar, comparar, entre otras, los posibles motivos que conllevan a la toma inadecuada de los datos, con el objetivo de lograr una mejora continua y que a futuro se posea un plan de contingencia frente a las variables intrínsecas y extrínsecas asociadas, que afecten la toma de datos por parte de las estaciones de control ambiental.





## Capítulo 3

### Marco Teórico

La constante evolución económica, industrial, tecnológica y comercial ha desembocado a un mundo más cómodo para el humano, pero al mismo tiempo se ha permitido germinar nuevos riesgos para la salud. Las investigaciones actuales aún carecen de suficiente soporte científico que permita identificar y aclarar con exactitud la relación que existe entre la salud y el medio ambiente. La falta de sistemas especializados que permitan la vigilancia sanitaria y la obtención de información que permita evaluar la gravedad y magnitud de los diferentes riesgos a los que continuamente estamos sometidos a nivel urbano lo hacen una actividad compleja. Se conoce que algunas sustancias presentes en el ambiente que están por debajo de ciertos límites no implican un peligro para la salud. Sin embargo, se encuentran contaminantes atmosféricos que suponen un riesgo en resultados inferiores a los límites inferiores permitidos en la actualidad [Vargas Marcos \(2005\)](#). El Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM) ha realizado esfuerzos por realizar un continuo seguimiento a los contaminantes ambientales que la Organización Mundial para la Salud recomienda mantener bajo vigilancia. En los últimos informes presentaron la compleja situación que viven los habitantes de algunas regiones y ciudades por los niveles elevados en los contaminantes atmosférico. Las Estaciones de monitoreo registraron concentraciones por encima de los niveles máximos establecidos. Desde el 2015 se vienen realizando declaraciones de sucesos donde los límites máximos de contaminantes atmosféricos superan de manera alarmante los niveles permitidos en la normatividad nacional, ocasionando un impacto en las actividades industriales, comerciales y de transporte entre otras. Se han buscado implementar políticas, acciones y estrategias que permitan la sensibilización y concientización de los ciudadanos. Sin embargo, las alertas son cada día más frecuentes en especial en época de pocas lluvias [IDEAM \(2018\)](#).

# Contaminación atmosférica

## Contaminación en el Aire

La calidad del aire en los centros urbanos, ha formado parte de los debates en calidad de vida y salud pública. Desde los años 90, se pronosticaba que la población mundial en incremento demandaría más vehículos de transporte. Generando una mayor cantidad de contaminación atmosférica por la combustión de estos vehículos [Fenger \(1999\)](#). Cuando se presentaron las primeras investigaciones sobre el tema, se tenía como contaminantes principales el hollín y el dióxido de azufre en las zonas urbanas. En la actualidad se estudian lo denominado como material particulado y gases producto de la combustión o gases que pueden generar algún tipo de daño en la salud que son producto de reacciones químicas en las que los gases de combustión favorecen la obtención. En donde se encuentran el material particulado menor a 10 micrómetros ( $PM_{10}$ ), material particulado menor a 2.5 micrómetros ( $PM_{2.5}$ ), dióxido de nitrógeno ( $NO_2$ ), dióxido de azufre ( $SO_2$ ), Monóxido de carbono (CO) y ozono troposférico ( $O_3$ ) [de Ambiente y Desarrollo Sostenible \(2017\)](#). Este último obtenido por la reacción del oxígeno, el dióxido de nitrógeno y luz ultra violeta. Lo anterior se debe principalmente a las relaciones que se tienen entre la calidad del aire y los indicadores de salud [Vecchi y cols. \(2008\)](#). De acuerdo a comunicados de la Organización Mundial de la Salud (OMS), en el año 2016 solo el 9 % de la población mundial vivía en zonas donde se tenían políticas y se seguía las recomendaciones de la OMS sobre calidad del aire. Desembocando en un número de defunciones prematuras de alrededor de 4,2 millones. Las enfermedades de respiratorias crónicas suman un total de muertes de 1.4 millones siendo la quinta causa de muerte derivadas del medio ambiente [de la Salud \(2016\)](#).

## Tipos de Contaminantes

Se pueden clasificar las fuentes de contaminación atmosférica según su fuente: Móviles: se asocia a la combustión interna de vehículos en general Fijas: Hace referencia a todas las industrias que emplean energía a partir de combustión de materia vegetal o combustibles fósiles para su funcionamiento. Incendios forestales: El cambio climático ha afectado parte de los bosques andinos volviéndolos sensibles a incendios forestales como resultado de la desecación [Young, Young, y Josse \(2011\)](#). Además, las fuentes de contaminantes también pueden clasificarse en dos grupos de acuerdo a su naturaleza: biodegradables y no biodegradables. Los biodegradables son materiales que pueden pasar por un proceso de descomposición por organismos como lombrices, hongos pluricelulares o microorganismos como las bacterias o levaduras, permaneciendo en el ciclo básico de nutrientes. De los que se pueden nombrar los derivados de material vegetativo o subproductos de la caña. Los no biodegradables no pueden ser descompuestos por la naturaleza o tardan prolongados

periodos de tiempo para lograrlo. Entre estos encontramos al petróleo, carbón mineral [Balcázar Gonzaga \(2013\)](#).

## Material Particulado

Es la mezcla de partículas en estado líquido y sólido de sustancias inorgánicas y orgánicas en suspensión, que varían en forma, composición y tamaño. La tabla **3-1** presenta los diferentes materiales particulados que la normatividad colombiana vigila según las recomendaciones de la OMS

Sustancia	Definición	Efectos sobre la salud
$PM_{10}$	Material Particulado menor a 10 micras ( $PM_{10}$ ): Son llamadas también partículas gruesas y tienen un tamaño comprendidos entre 2.5 y 10 micrómetros. Aproximadamente el ancho de un séptimo de cabello.	Afectan el tracto respiratorio superior por lo que se conocen como fracción torácica. Ocasiona irritación aguda de los ojos y de las vías respiratorias. Reducción de la capacidad pulmonar y agotamiento respiratorio. Desarrollo de problemas respiratorios y cardiovasculares.
$PM_{2.5}$	Material Particulado menor a 2.5 micras ( $PM_{10}$ ): Se conocen como partículas finas de menor de 2.5 micras de diámetros. Suponen un mayor peligro para la salud que el PM10, debido a que al inhalarlas pueden alcanzar zonas periféricas de los bronquiolos.	También denominada fracción respirable. Influye en el desarrollo de cardiopatías y neuropatías, así como enfermedades crónicas de obstrucción pulmonar. Agravan el asma y se asocian al desarrollo de diabetes, bajo peso al nacer e influye en una disminución en el desarrollo cognitivo de los niños.

Tabla **3-1**: Contaminantes atmosféricos particulados. Fuente: [IDEAM \(2018\)](#)

## Gases

Los gases que hacen parte de los contaminantes atmosféricos son sustancias y compuestos en estado gaseoso que influyen o alteran la química atmosférica y que afectan directamente la salud. La tabla **3-2** nos muestra las sustancias y el efecto sobre la salud.

Sustancia	Definición	Efectos sobre la salud
Dióxido de Nitrógeno ( $NO_2$ ):		
$NO_2$	En un proceso paralelo al $SO_2$ , el nitrógeno en los combustibles se convierte por combustión a altas temperaturas a óxidos de nitrógeno, $NO_x$ , que corresponden a la suma de $NO_2$ y NO	Irrita las vías respiratorias. Causa bronquitis y pulmonía. Reduce significativamente la resistencia respiratoria a las infecciones
Dióxido de Azufre ( $SO_2$ ):		
$SO_2$	Gas incoloro que se forma a partir de la combustión de sustancias que contienen azufre, principalmente petróleo y carbón, así como de numerosos procesos industriales	Causa bronco constricción, bronquitis y traqueítis. Agrava enfermedades respiratorias y cardiovasculares existentes
Ozono troposférico ( $O_3$ ):		
$O_3$	Es un gas que no se emite directamente por fuentes primarias. Se produce a partir de reacciones fotoquímicas en presencia de radiación solar y precursores, tales como los óxidos de nitrógeno ( $NO_x$ ) y los compuestos orgánicos volátiles (COV).	Irrita el sistema respiratorio. Reduce la función pulmonar. Agrava el asma. Inflama y daña las células que recubren los pulmones. Agrava las enfermedades pulmonares crónicas. Causa daño pulmonar permanente
Monóxido de Carbono (CO):		
CO	Se forma a partir de la combustión incompleta de combustibles que contienen carbono. Este es un caso común donde una proporción del carbón se oxida solamente a monóxido de carbono, mientras que la combustión completa conduce a la formación de dióxido de carbono.	Inhabilita el transporte de oxígeno a las células. Provoca mareos, dolor de cabeza, náuseas, estados de inconsciencia e incluso la muerte.

Tabla 3-2: Contaminantes atmosféricos gaseosos. Fuente: [IDEAM \(2018\)](#)

## Ozono Troposférico

Cuando miramos al cielo encontramos un característico color azul. Este color es originado por la presencia de ozono el cual es reconocido el constituyente de la capa protectora de nuestro planeta. Sin embargo, el tenerlo presente en las capas atmosféricas en la que vivimos resulta perjudicial para nuestra salud y el medio ambiente [Phillips \(2000\)](#). Cuando se habla de ozono de baja altura se hace referencia al ozono de la tropósfera. Además, posee complejas y dinámicas reacciones químicas que involucra a otras sustancias químicas

y condiciones atmosféricas. Las reacciones fotoquímicas formadoras de ozono troposférico involucran óxidos de nitrógeno ( $NO_x$ ) y compuestos orgánicos volátiles (COV). Estos últimos son provenientes de las emisiones de vehículos de combustión interna, fabricas e incendios de cualquier tipo [Borsese y Esteban \(2005\)](#). De acuerdo a la Resolución 2254 de 2017, el límite máximo es de  $100 \mu g/m^3$  con una exposición de 8 horas. Mientras que la Resolución 610 de 2010, establece los límites máximos de  $O_3$  en  $120 \mu g/m^3$  por un periodo de 1 hora y  $80 \mu g/m^3$  para un tiempos de exposición de 8 horas. Par este contaminante no existe un tiempo de exposición anual, por lo que no se cuenta con un valor de referencia o comparativo en ese intervalo de tiempo [de Ambiente y Desarrollo Sostenible \(2017\)](#)

## Normatividad Colombiana Aplicable

Los niveles máximos permitidos para los contaminantes atmosféricos nacionales, fueron establecidos en primera instancia por la Resolución 610 de 2010 y renovados por la Resolución 2254 de 2017. Tanto la resolución 610 y 2254 adoptan las sugerencias suministradas por la OMS, la cual, establece rangos de acuerdo a estudios epidemiológicos de mortalidad y movilidad que relacionan enfermedades cardiovasculares y respiratorias que se asocian directamente con contaminantes atmosféricos [IDEAM \(2018\)](#). A continuación, se presenta la tabla **3-3** se presentan los diferentes contaminantes y la normatividad que establece los valores permitidos por la legislación colombiana y por la OMS.

Contaminante	Res. 610 de 2010	Res. 2254 de 2017 (a partir de 2018)	Res. 2254 a 2030 (a partir de 2030)	OMS Objetivo Intermedio 2	OMS Objetivo Intermedio 3	Valor Guía OMS
<b>Valores anuales – Exposición prolongada</b>						
$PM_{10}$	50	50	30	50	30	20
$PM_{2.5}$	25	25	15	25	15	10
$NO_2$	100	60	40	-	-	40
<b>Valores diarios – Exposición de corta duración – 24 horas</b>						
$PM_{10}$	100	75	75	100	75	50
$PM_{2.5}$	50	37	37	50	37.5	25
$SO_2$	250	50	20	-	-	-
<b>Valores octohorarios – Exposición de corta duración – 8 horas</b>						
$O_3$	80	100	-	-	-	100
CO	10000	5000	-	-	-	-

Tabla **3-3**: Normatividad de calidad del aire en Colombia vrs. Recomendaciones de la Organización Mundial de la Salud. Las unidades correspondientes a cada uno de los contaminantes son, para  $PM_{10}$   $\mu g/m^3$ ,  $PM_{2.5}$   $\mu g/m^3$ ,  $NO_2$  ppm,  $SO_2$  ppm,  $O_3$  ppm, CO ppm Fuente: [IDEAM \(2018\)](#)

## Índice de Calidad del Aire – ICA

El IDEAM calcula un valor adimensional y asigna una categoría y un color para determinado rango de concentración de los contaminantes atmosféricos de acuerdo al daño que pueden causar a la salud de las personas.

ICA	COLOR	CLASIFICACIÓN	$PM_{10}$ $\mu g/m^3$	$PM_{2.5}$ $\mu g/m^3$	CO ppm
0 - 50	Verde	Buena	0 - 54	0.0 - 15.4	0.0 - 4.4
51 - 100	Amarillo	Moderada	55 - 154	15.5 - 40.4	4.5 - 9.4
101 - 150	Naranja	Dañina a la salud para grupos sensibles	155 - 254	40.5 - 65.4	9.5 - 12.4
151 - 200	Rojo	Dañina a la salud	255 - 354	65.5 - 150.4	12.5 - 15.4
201 - 300	Púrpura	Muy dañina a la salud	355 - 424	150.5 - 250.4	15.5 - 30.4
301 - 400	Marrón	Peligrosa	425 - 504	250.5 - 350.4	30.5 - 40.4
401 - 500	Marrón	Peligrosa	505 - 604	350.5 - 500.4	40.5 - 50.4

Tabla **3-4**: Puntos de corte del Índice de Calidad del Aire ICA en Colombia para material particulado y dióxido de carbono. Fuente: [IDEAM \(2018\)](#)

Este valor recibe el nombre de Índice de Calidad del Aire (ICA). El ICA, es calculado para seis de los siete contaminantes contemplados en la normativa colombiana ( $PM_{10}$ ,  $PM_{2.5}$ , CO,  $O_3$ ,  $SO_2$  y  $NO_2$ ) y tiempos de exposición (1 hora y 24 horas) de acuerdo con los parámetros establecidos en el Protocolo para el Monitoreo y Seguimiento de la Calidad del Aire. La escala ICA, se encuentra entre 0 y 500, representando 500 el mayor nivel de contaminación atmosférica y, por consiguiente, mayor daño a la salud. Este indicador busca presentar de una manera fácil y eficaz la condición del aire y su efecto en la salud a todos los habitantes que deben vivir en medio de los siete contaminantes. En la tabla **3-4** y **3-5**, se presenta los siete contaminantes objeto del monitoreo y seguimiento a excepción de partículas suspendidas totales. Como base, fue utilizada la Resolución 610 de 2010

ICA	COLOR	CLASIFICACIÓN	O <sub>3</sub> ppm	O <sub>3</sub> ppm	SO <sub>2</sub> ppm	NO <sub>2</sub> ppm
0 - 50	Verde	Buena	0.000 -	-	0.000 - 0.034	(2)
51 - 100	Amarillo	Moderada	0.059 0.060 -	-	0.035 -	(2)
101 - 150	Naranja	Dañina a la salud para grupos sensibles	0.075 0.076 -	0.125 -	0.144 0.145 -	(2)
151 - 200	Rojo	Dañina a la salud	0.095 0.096 -	0.164 0.165 -	0.224 0.225 -	(2)
201 - 300	Púrpura	Muy dañina a la salud	0.115 0.116 -	0.204 -	0.304 0.205 -	0.65 -
			0,374 (0,155 - 0,404) (4)	0.404	0.604	1.24
301 - 400	Marrón	Peligrosa	(3)	0.405 -	0.605 -	1.25 -
				0.504 0.505	0.804 0.805	1.64 1.65
401 - 500	Marrón	Peligrosa	(3)	-	-	-
				0.604	1.004	2.04

Tabla **3-5**: Puntos de corte del Índice de Calidad del Aire ICA en Colombia para gases atmosféricos contaminantes. Fuente: [IDEAM \(2018\)](#)

(2) Para NO<sub>2</sub> se tendrán en cuenta valores únicamente por encima de 200 teniendo en cuenta que han sido tomados de valores y parámetros EPA. (3) Valores de concentraciones de 8 horas de ozono no definen valores más altos de ICA ( $\geq 301$ ). Los valores de ICA de 301 o mayores serán calculados con concentraciones de 1 hora de ozono. (4) Los números entre paréntesis se asocian a valores de 1 hora que se utilizarán en esta categoría sólo si se superponen.

## Estadística aplicada

La estadística aplicada se encarga de conocer los resultados o características de una población a partir de una muestra o el estudio de un censo. Para realizar ese estudio, la estadística emplea datos para ahondar en temas y llegar a una conclusión de interés general o particular. Herramientas gráficas y análisis en base a cálculos matemáticos, hacen parte de su esencia [Moore \(2005\)](#). Se observan datos y se determinan medidas de posición central y no central, correlaciones, probabilidades, Índices, entre muchos más, de carácter invariado o multivariado. A continuación, se presenta algunas herramientas estadísticas



que se emplearán en este trabajo de grado.

## Muestreo Probabilístico

Entre las preguntas más frecuentes y generadoras de conflictos es el identificar cuánto es la población y bajo qué argumentos será seleccionada la muestra, su cantidad y demás componentes que la conforman. Estas preguntas se hacen en diferentes ramas psicología, historia, sociología, biología, medicina, ingeniería, etc. [López \(2004\)](#)

### Población

Hace referencia al conjunto de objetos o personas de interés para el desarrollo de la investigación. " *El universo o población puede estar constituido por personas, animales, registros médicos, los nacimientos, las muestras de laboratorio, los accidentes viales entre otros*". [Canales, Alvarado, y Pineda \(1994\)](#).

### Muestra

Es un grupo de una población o universo del que se obtienen datos de manera representativa. [Hernández-Sampieri, Fernández-Collado, y Baptista-Lucio \(2017\)](#)

### Muestreo

Es la manera o método que se encargara de seleccionar la cantidad de individuos o datos a estudiar. Se asocia a reglas, criterios y procedimientos necesarios para garantizar que sea representativo de la población o universo. [Mata \(1994\)](#)

### Tamaño de la Muestra

Existe diferentes ecuaciones matemáticas que permiten calcular el tamaño de una muestra. Sin embargo, se deben tener en cuenta diferentes criterios para elegir la correcta. Entre ellos está si la muestra es infinita o finita, si es en base en proporciones utilizada con poblaciones que las variables son de carácter categórico o numérico, entre otros. Pero es muy importante conocer los recursos económicos y que se requiere por parte del investigador. También se debe considerar que tan grande es el margen de error que está dispuesto a tolerar, como también el nivel de confianza deseado, que por lo general es del 95 % [Hernández-Sampieri y cols. \(2017\)](#).

Otro elemento a considerar para el tamaño de la muestra, son ecuaciones matemáticas de las cuales sólo se presentará una ecuación simple y fácil de aplicar (3-1). Es de mencionar, que las ecuaciones estadísticas se basan en el margen de error, la confiabilidad la probabilidad.

$$n = \frac{N}{((N - 1) \times K^2) + 1} \quad (3-1)$$

Donde  $n$  es el tamaño de muestra,  $N$  el tamaño de la población y  $K$  es el margen de error.

### Tamaño de la Muestra según la proporción

El tamaño de muestra depende directamente de la varianza ( $\sigma$ ) poblacional 3-2. La cual, puede ser determinada por datos históricos. Sin embargo, en muchas ocasiones no se cuenta con dicha información. Por lo cual se procede a realizar cálculos que involucren la proporción  $p$ . Por lo general, los valores de la proporción ( $p$ ) poblacional y su complemento ( $q=1-p$ ) también son desconocidos. Pero al realizar una prueba donde se asignan valores a  $p$  y  $q$ , el mayor valor a tomar sería de ( $p=q=0.5$ ). Como resultado se obtiene el máximo valor que resulta para el tamaño de la muestra dependiendo del margen de error y la confianza [Morillas \(2007\)](#).

$$n = \frac{Z_{1-\frac{\alpha}{2}}^2 \times \sigma}{\varepsilon^2} \quad (3-2)$$

El tamaño de muestra ( $n$ ) es determinada por la siguiente ecuación 3-3:

$$n = \frac{Z_{1-\frac{\alpha}{2}}^2 \times p \times q}{\varepsilon^2} \quad (3-3)$$

Donde  $\varepsilon$  es el margen de error.

### Tipo de Muestreo

Como se mencionó con anterioridad, existe dos tipos, muestreo probabilístico y no probabilístico. Debido que este trabajo busca garantizar la misma probabilidad se enfocará en el muestreo probabilístico. El muestreo probabilístico tiene cuatro diferentes métodos los cuales son aleatorio, sistemático, estratificado y por conglomerado, de los cuales solo se hará énfasis en el estratificado y sistemático pues son métodos utilizados en este trabajo de grado.

#### Muestreo Probabilístico Sistemático

Este método busca establecer un intervalo que se respetará para la toma de las observaciones. Algunos investigadores consideran que esta técnica es relevante para investigaciones sociales de gran tamaño, como lo señala [López \(2004\)](#) y que lo referencio de Torres B (1997) Este muestreo es empleado tamaños de poblaciones grandes, pero también en pequeñas donde el método facilita la caracterización de los componentes de la muestra Este tipo de muestreo se utiliza cuando el tamaño de la población es grande y la tabla de números aleatorios no es suficiente para contar esa población. También es usado en poblaciones pequeñas donde la selección sistemática facilita la identificación de los componentes de la muestra. [López \(2004\)](#).

## Análisis Multivariado

Es común ver que muchos de las investigaciones o trabajos de grado de pregrado se encuentren tratamientos de datos con el uso de estadística descriptiva univariada. El utilizar una tercera variable que se asocie con las dos primeras, adquiere un nuevo más complejo grado de dificultad. Este grado es proporcional al número de variables que arroja el trabajo de investigación o de grado. Los cálculos asociados llevan inmersos el tratamiento de una cantidad de datos elevada de las variables convirtiéndose en una de las principales limitaciones por la que el investigador solo hasta hace algunos años ha visto la necesidad de implementar un tipo de análisis multivariado [Lozares Colina y López-Roldán \(1991\)](#). Las ventajas que presenta el análisis multivariado sobre el univariado o bivariado son las siguientes [Nesselroade y Cattell \(2013\)](#):

- Menor necesidad de almacenamiento de los datos
- Aumento de la seguridad en la inferencia estadística,
- Mejora de conceptos teóricos más pertinente,
- Incremento de la perspectiva y la precisión de conceptos.

## Análisis de Componentes Principales

En la estadística multivariada existe una herramienta de gran utilidad y objetividad para el manejo de datos y posterior evaluación obtenidos por medio del Sistemas de Vigilancia de Calidad del Aire denominado análisis de componentes principales (ACP) [Martínez \(2016\)](#) o técnicas de matemáticas de clasificación como el análisis de conglomerados. No solo aplicando a la calidad del aire, sino también a la evaluación y manejo de seguimientos de calidad del agua [Shrestha y Kazama \(2007\)](#), [Singh, Malik, Mohan, y Sinha \(2004\)](#), [Alberto y cols. \(2001\)](#). El uso combinado de estas dos técnicas de ha llevado a cabo para determinar distribución espacial y tendencias estacional de  $O_3$  y  $PM_{10}$  [Gramsch, Cereceda-Balic, Oyola, y Von Baer \(2006\)](#). El uso del ACP en el Sistemas de Vigilancia de Calidad del Aire, ha promovido su aplicación de manera frecuente en diferentes lugares del planeta [Ibarra-Berastegi y cols. \(2008\)](#), [Ibarra-Berastegi y cols. \(2009\)](#), [Pires, Sousa, Pereira, Alvim-Ferraz, y Martins \(2008\)](#), [Pires, Pereira, Alvim-Ferraz, y Martins \(2009\)](#).

## Control Estadístico de la Calidad

Durante los diferentes procesos industrializados o administrativos, se hace necesario realizar seguimientos que permitan disminuir pérdidas por insatisfacción de los clientes o pérdida de materia prima. Por lo que se hace necesario realizar un control sobre las diferentes variables que se asocian a un proceso. Una de las definiciones más precisas en relación a la calidad y la variabilidad, es la relación inversamente proporcional que existe

entre estas dos [Montgomery \(2012\)](#). Para determinar la variabilidad, es necesario utilizar herramientas que permitan cuantificar la variación o calificarla en un proceso, producto o servicio para analizar los resultados y de esa manera tomar las mejores decisiones y como resultado reducir o controlar la variabilidad exigida para cumplir con los requisitos establecidos [Velásquez Martínez y cols. \(2019\)](#). El control estadístico de calidad, emplea múltiples herramientas de análisis técnico, como son regresiones lineales múltiples, diagramas y gráficos de control que pueden llegar hacer generales de un proceso o con énfasis en pequeñas desviaciones de las variables del proceso.

### **Gráficos de control**

Los gráficos de control de calidad, son instrumentos estadísticos diseñados para representar visualmente la variabilidad de las particularidades de un producto, proceso o servicio para analizar con mayor facilidad que la ofrecida por una tabla con observaciones o resultados de una serie de cálculos. Los gráficos representan las observaciones o mediciones que por lo general es por intervalos de tiempos previamente definidos y fijos. Y los límites de control superior, inferior y la media de estos, según sea el caso [Velásquez Martínez y cols. \(2019\)](#). Entre las ventajas que ofrece los gráficos de control de calidad encontramos [Ruiz Barzola \(2013\)](#):

- Visualizar e identificar comportamientos estadísticos.
- Reconocer causas que intervienen en el proceso, producto o servicio.
- Evaluar la capacidad de un proceso según los requisitos iniciales establecidos, reconociendo las ocasiones de mejora.



# Capítulo 4

## Diseño Metodológico

### Método de Estudio

#### Método

Se desarrolló una investigación descriptiva, partiendo del análisis univariado descriptivo y un posterior multivariado, para buscar establecer características que afecten directamente una variable, ya sea incrementándola o disminuyéndola de acuerdo a variables de carácter natural.

#### Enfoque

Enfoque cuantitativo: Partiendo de datos cuantitativos se realizará una discusión de los resultados alcanzados del análisis estadístico, esto, con la intención de identificar posibles alteraciones que pueda sufrir el ozono en diferentes condiciones climatológicas.

### Diseño Metodológico

Este trabajo de grado se desarrollará en tres principales fases: Obtención y organización de la fuente de datos, métodos estadísticos descriptivos (univariado y multivariado) y cartas de control estadístico. Y como última fase, el análisis de resultados asociando el ozono a su comportamiento con las condiciones climatológicas y su cumplimiento con la normatividad colombiana.

#### Fase 1. Fuente y organización de la información

El gobierno nacional colombiano por medio de la página web Datos Abiertos, administrada por el Ministerio de Tecnologías de la Información y las Comunicaciones, deja a disposición del interesado bases de datos de entidades públicas con múltiples temas de interés nacional, regional o local. A continuación, se presentará los datos relacionados a la base de datos seleccionada.

### ***Base de datos***

- Nombre: Datos de calidad del aire en la estación Catumare Junio a Diciembre 2018
- Fuente: Corporación para el Desarrollo Sostenible del Área de Manejo Especial la Macarena
- Descripción: La estación CATUMARE mide los niveles de Ozono (O<sub>3</sub>) y variables meteorológicas como velocidad del viento, dirección del viento, temperatura, humedad, radiación solar y precipitación.
- Archivo: Datos\_de\_calidad\_del\_aire\_en\_la\_estaci\_n\_Catumare\_Junio\_a\_Diciembre\_2018.csv
- link pagina web:<https://www.datos.gov.co/Agricultura-y-Desarrollo-Rural/Datos-de-calidad-del-aire-en-la-estaci-n-Catumare-/3xvf-rxsa?defaultRender=table>

### **Datos Empleados para el estudio**

- Proyecto: Política de Prevención y Control de la Contaminación del Aire
- Zona de Estudio: Estación ambiental Catumare
- Fecha de muestreo: 25 de junio al 31 de diciembre de 2018
- Técnica de medición: Unidades sistematizadas autónomas que implementan sensores electrónicos. Se desconoce si es certificado o no el funcionamiento de los sensores.
- Tipo de muestra: Líquida y gaseosa

### **Fase 2. Análisis Estadístico**

#### **Organización de la base de datos**

La base de datos se organizó de acuerdo al promedio semanal para cada una de las variables exonerando los valores iguales a cero para la variable ozono. Si el equipo de toma de datos realiza al menos un solo registro, se tuvo en cuenta como promedio semanal. Al no presentar al menos uno de los 168 registros posibles semanales, se descartó la correspondiente semana.

#### **Análisis Descriptivo**

Se efectuó un análisis descriptivo univariado buscando una caracterización de la observación mínima y máxima, promedio, desviación estándar y prueba de Shapiro como evaluación del comportamiento normal de las observaciones.

## **Análisis Multivariado**

Buscando reducir de ocho a menos variables como paso preliminar al control estadístico de la calidad, se realizó un análisis de componentes principales (ACP). Además, identificar las variables meteorológicas que no afectan o afectan de manera proporcional e inversamente proporcional al ozono.

## **Cartas de Control Estadístico**

Empleando cartas de control estadístico multivariado, se espera encontrar posibles observaciones que se encuentren fuera de los límites de control, evaluando los componentes principales, buscando explicar la mayor variabilidad de los datos.

## **Fase 3. Análisis de Resultados**

Partiendo de la importancia que tiene el realizar una adecuada toma de datos y aún más su correcto análisis frente a las condiciones meteorológicas, se tendrá en cuenta las diferencias de un análisis univariado frente al multivariado y la relación que puede existir con el resto de las variables.





## Capítulo 5

# Análisis y Resultados

La base de datos está compuesta por 4455 observaciones tomadas cada hora desde las 0:00 horas del 25 de junio hasta las 13:59 horas del 31 de diciembre de 2018 que comprende 27 semanas. De las ocho variables, se tiene el ozono ( $\mu g/m^3$ ) como único contaminante atmosférico y Dirección del viento (Grados Sexagesimales), humedad (%), precipitación ( $l/m^2$ ), presión atmosférica (mbar), Radiación solar ( $W/m^2$ ), temperatura ( $^{\circ}C$ ) y velocidad del viento (m/s) como condiciones meteorológicas. Un estudio univariado preliminar de la base de datos evidenció que alrededor del 36.99 % de los datos de ozono corresponden al valor cero (0). Teniendo en cuenta que en la composición del aire encontramos ozono y que se contaba con registros previos de una resolución de 0.01, es posible que la estación ambiental sufriera algún tipo de avería que no permitía registrar el dato de ozono, pero si los datos de las variables meteorológicas.

### Organización y Toma de Muestra

Debido a la presencia de valores en cero y con la intención de evaluar los datos suministrados por la estación ambiental, se realizó una depuración en la cual se excluyen la fila de registros (esto incluye el contaminante atmosférico y las siete condiciones atmosféricas) en donde las observaciones de ozono sean iguales a cero. Esto permitió garantizar que las otras variables no presentaran observaciones en cero a acepción de las variables precipitación y radiación solar. En total se inició con observaciones de 27 semanas. Debido a la cantidad de datos presentes, se decide realizar el muestreo aleatorio por semana para realizar los diferentes análisis que se someterá la base de datos. Pero teniendo en cuenta que no todas las semanas tienen la misma cantidad de observaciones de ozono, se realizó la siguiente figura.

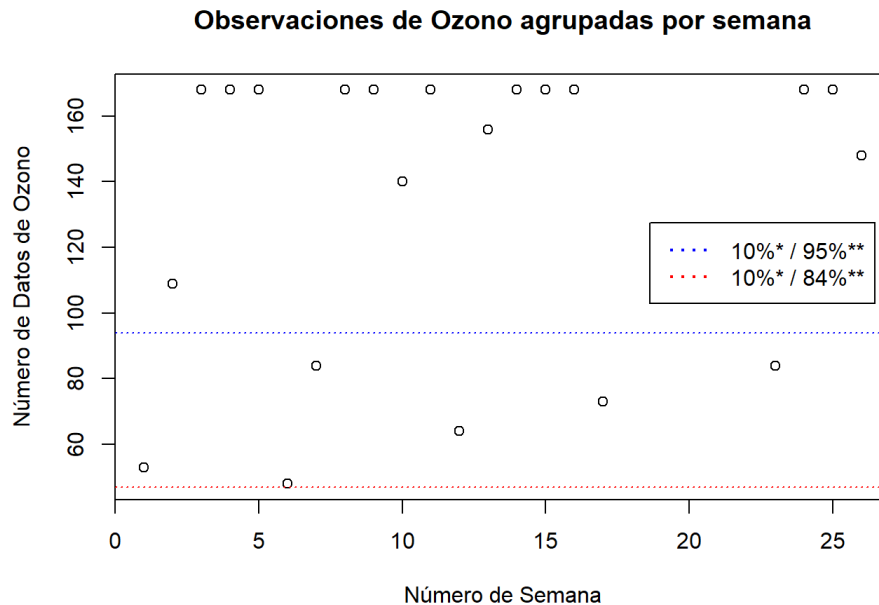


Figura 5-1: Selección de la cantidad de muestra a utilizar por semana. La mínima cantidad de observaciones por muestra se da en el día seis con un total de 48. \*Margen de error y \*\*Nivel de confianza para una proporción de éxito ( $p = 0.5$ ) y de fracaso ( $q = 0.5$ ) como se observa en la ecuación 3-3

En la figura 5-1 se ilustra la cantidad de observaciones de ozono para las 27 semana donde se presenta un periodo aproximado de seis semanas sin registro (semanas 18, 19, 20, 21, 22 y 27), 10 semanas con registros incompletos y 11 semanas con registros completos. Entre las semanas se tiene como máximo 168 observaciones y como mínimo 48. Buscando hacer un muestreo aleatorio se realiza los cálculos con un margen de error del 10 % y nivel de confianza del 95 %, pero la cantidad de muestras es de 92 para una población finita y 96 para una infinita. Debido a que se buscó tomar una muestra equitativa entre las semanas, se decide tomar el menor valor de observaciones de ozono que para este caso corresponde a la semana seis con 48 observaciones. En la figura 5-1 se ilustra que el tomar una muestra de 48 observaciones por semana, equivale a tener un margen de error del 10 % con un nivel de confianza del 84 % para una probabilidad de éxito del 50 %. A cada una de las ocho variables se le obtuvo el promedio de sus 48 datos por semana, resultando una base de datos a la cual se le aplicó el estudio estadístico.

## Análisis Descriptivo Univariado

Se realizó cálculos correspondientes a medidas de posición central, dispersión, forma, test de distribución normal univariada y un análisis específico para la variable ozono como se presenta a continuación.

## Medidas de tendencia central, dispersión y forma

Se realizó el análisis descriptivo estándar con el fin de evidenciar la dispersión que existe entre los datos y la media de cada variable y su correspondiente distribución de frecuencia.

	$O_3$ ( $\mu g/m^3$ )	Dirección del Viento ( $^\circ$ )	Humedad (%)	Precipitación ( $l/m^2$ )
<b>Mínimo</b>	12.176	158.688	65.979	0.021
$Q_1$	19.826	-	72.292	0.085
<b>Promedio</b>	23.368	-	76.007	0.219
<b>Mediana</b>	22.966	-	75.062	0.185
$Q_3$	28.079	-	79.594	0.320
<b>Máximo</b>	34.656	231.708	86.812	0.485
<b>Varianza</b>	30.015	-	24.513	0.018
<b>Desviación Estandar</b>	5.479	-	4.951	0.135
<b>CV( %)</b>	23.447	-	6.514	61.644
<b>Coefficiente de Curtosis</b>	-0.674	-	-0.465	-1.151
<b>Coefficiente de Asimetría</b>	-0.065	-	0.169	0.187
	Presión Atmosférica (mbar)	Radiación Solar ( $W/m^2$ )	Temperatura ( $^\circ C$ )	Velocidad del Viento (m/s)
<b>Mínimo</b>	959.917	46.604	22.544	0.642
$Q_1$	961.375	96.073	23.802	0.829
<b>Promedio</b>	962.440	119.717	24.642	0.895
<b>Mediana</b>	962.854	121.104	24.883	0.887
$Q_3$	963.427	146.010	25.311	0.971
<b>Máximo</b>	964.854	190.771	27.017	1.119
<b>Varianza</b>	2.008	1242.606	1.208	0.012
<b>Desviación Estandar</b>	1.417	35.251	1.099	0.110
<b>CV( %)</b>	0.147	29.445	4.460	12.290
<b>Coefficiente de Curtosis</b>	-1.057	-0.500	-0.622	-0.254
<b>Coefficiente de Asimetría</b>	-0.334	-0.172	0.002	-0.071

Tabla 5-1: Análisis descriptivo univariado del promedio del muestreo de 21 semanas

\* Rango Inter Cuartílico.

De acuerdo a la tabla 5-1, el comportamiento que tienen los datos con relación a la media, permite identificar tres grupos de acuerdo al coeficiente de variación (CV). Entre las variables con menor variación en relación a la media se encuentra la presión atmosférica, la

temperatura y la humedad. Mientras que la dirección y velocidad del viento se encuentran entre el nueve y 12 %. Además, se encuentran la radiación solar y el ozono con porcentaje similares y la precipitación con el mayor CV. Lo que indica los diversos cambios climáticos en los que está expuesta la estación ambiental. En especial con relación a lluvias y nubosidad. Tanto la humedad y la precipitación presentan una distribución asimétrica positiva, mientras que la presión atmosférica y la radiación solar presentan una asimetría negativa. En el caso del ozono y la velocidad del viento son asimétricas negativas, pero en menor proporción que las dos anteriores. Todas las variables presentan una distribución platicúrtica, indicando una reducida acumulación de datos en la posición central.

## Test de distribución normal univariado

Debido a que se cuenta con 21 observaciones, se hace necesario la aplicación del Test de Normalidad de Shapiro Wilk. Según la tabla **5-2**, todas las variables tienen un p-valor mayor a 0.05. Por lo tanto, existe suficiente evidencia para que la hipótesis nula, que indica que las observaciones provienen de una distribución normal, no es rechazada.

	Ozono	Dirección del Viento	Humedad	Precipitación
<b>W</b>	0.9853	0.9842	0.9909	0.9590
<b>p-valor</b>	0.9806	0.9726	0.9988	0.4970
	Presión Atmosferica	Radiación Solar	Temperatura	Velocidad del Viento
<b>W</b>	0.9485	0.9852	0.9783	0.9884
<b>p-valor</b>	0.3193	0.9798	0.8999	0.9945

Tabla **5-2**: Resultado por variable del test de normalidad de Shapiro Wilk

En el caso de la variable ozono, se puede identificar la distribución normal que tiene y la ausencia de datos atípicos como en la figura **5-2** se ilustra. En el diagrama de densidad del ozono, se puede identificar la asimetría negativa acumulando más datos a la derecha de la media que a la izquierda. Mientras que la gráfica de Distribución normal del ozono, reafirma el resultado de la prueba de Shapiro Wilh.

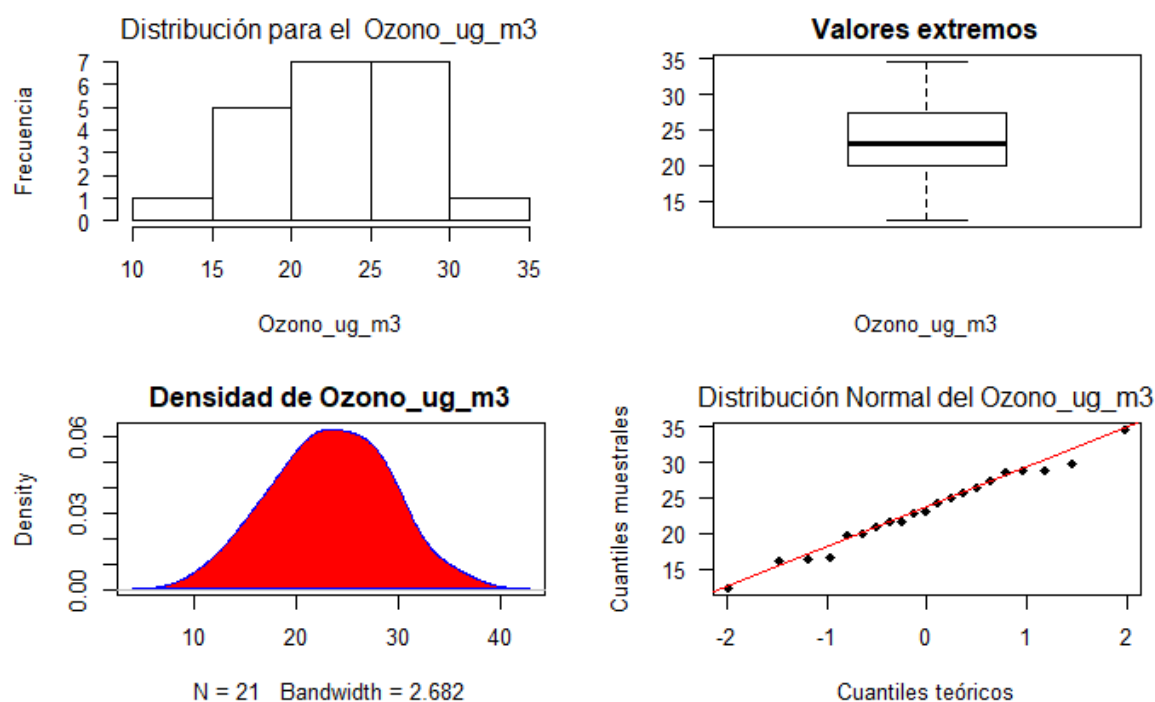


Figura 5-2: Análisis univariado para el contaminante atmosférico Ozono

## Análisis Descriptivo Multivariado

### Análisis de Correlación

En la figura 5-3 se identifican las variables meteorológicas que se relacionan directamente con la reacción fotoquímica necesaria en la formación de ozono. Entre estas encontramos la temperatura con una fuerte correlación positiva mientras que la velocidad del viento y la radiación solar es más débil. La humedad y la presión atmosférica evidencian una correlación negativa con el ozono fuerte. La presión atmosférica es uno de los factores que afecta directamente la humedad relativa y la capacidad del aire de retener agua. [Medrano \(2003\)](#). Tanto la humedad, como la temperatura son variables que tienen una fuerte correlación negativa. La humedad en el ambiente refleja días poco soleados y al ser una región con abundante Zonas selváticas, hacen de Villavicencio una ciudad húmeda por naturaleza. Sin embargo, las altas temperaturas se asocian con días despejados, el ingreso de rayos ultravioletas necesarios para la formación de ozono troposférico. La velocidad del viento no es un factor de que promueva la formación de Ozono. Sin embargo, actúa como medio de transporte para los contaminantes desde los lugares de generación de ozono y las unidades ambientales [IDEAM \(2018\)](#).

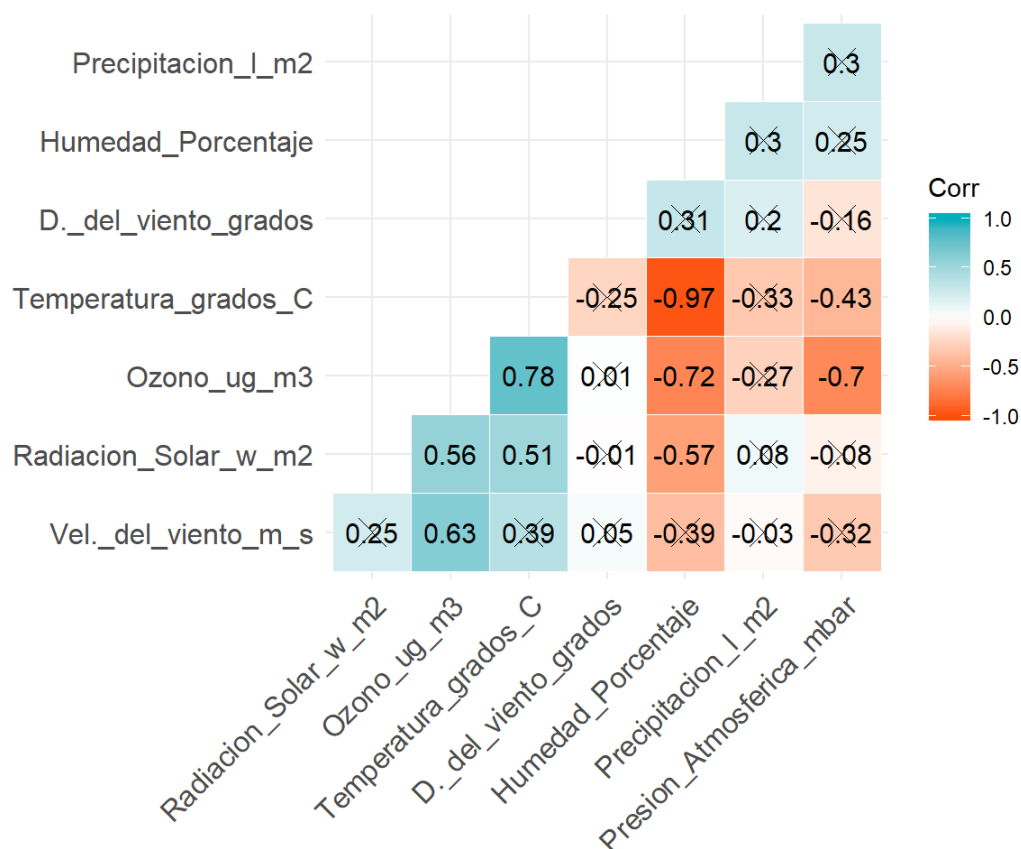


Figura 5-3: Matriz de correlación de Pearson

Entre las siete variables meteorológicas asociadas al ozono solo dos variables, la precipitación y en mayor medida la dirección del viento, no presentan correlación con el contaminante ozono. Al no ser parte de los diferentes precursores o mitigadores del contaminante ozono, como si lo hace en óxidos y material particulado, se retira de la base de datos para centrarnos en las condiciones meteorológicas que presentaron correlaciones fuertes o débiles con el ozono. Por esta razón se decidió trabajar en los análisis estadísticos, de aquí en adelante, con las variables Humedad, Presión atmosférica, Radiación solar, Temperatura y Velocidad del viento.

### Test de distribución Normal Multivariado

Se presentan los diferentes métodos por los cuales se busca identificar y justificar el tipo de distribución multivariada que presentan los datos, como argumento para la implementación de un análisis de componentes principales y el estudio de control de calidad en los datos. Para este fin se someterán al método gráfico y se confirmara por medio del método sesgo de Mardia.

## Método Gráfico: Distancia generalizada

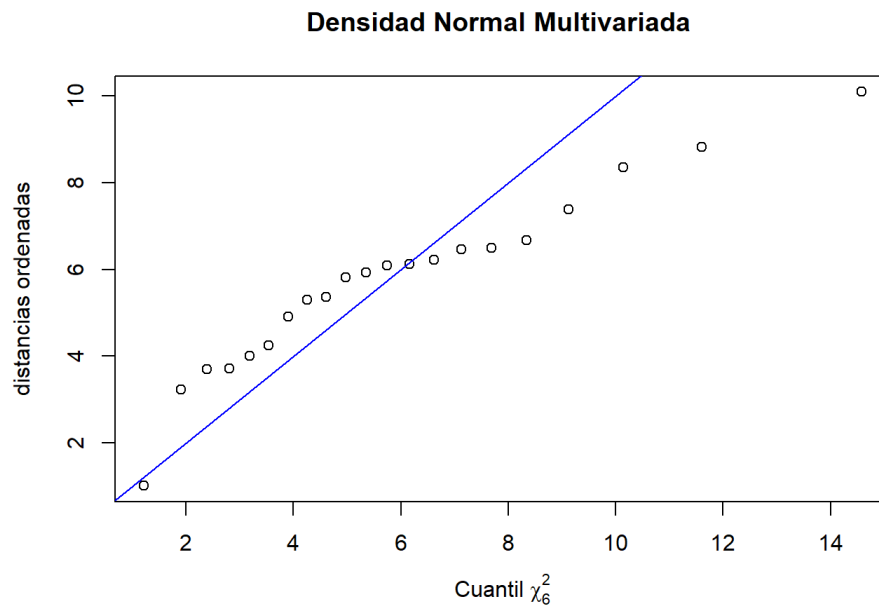


Figura 5-4: Determinación gráfica de distribución normal multivariable por medio de la distribución chi-cuadrado con seis grados de libertad

En la figura 5-4 busca identificar si las observaciones proviene de una distribución chi cuadrado con seis grados de libertad, que está representada por la línea azul. Sin embargo, para este caso no resulta fácil identificar dicho comportamiento en los datos ya que se espera que los puntos estén sobre la línea azul. Dejando al método sesgo de Mardia que confirme si es o no una distribución normal multivariada. Los datos no presentaron datos atípicos al comparar el valor de la distancia generalizada con el valor de chi cuadrado 0.05 a la derecha.

## Test de Multinormalidad: Prueba de Mardia

Test	Estadística	p-Valor
Asimetría de Mardia	30.8868	0.9974
Curtosis de Mardia	-1.7789	0.07524

Tabla 5-3: Prueba de Mardia para determinar multinormalidad



La tabla **5-3** presentan los resultados después de aplicar la Prueba de Mardia, la cual realiza prueba de hipótesis en dos test diferentes. Para afirmar que los datos tienen una distribución normal multivariada se necesita que los valores del p-valor sea superior a 0.05 par no rechazar la hipótesis nula. De esa manera, existe suficiente evidencia para no rechazar la hipótesis nula. Por lo tanto, los datos tienen una distribución normal multivariada y facilitará obtener, de manera adicional, algunos resultados inferenciales adicionales.

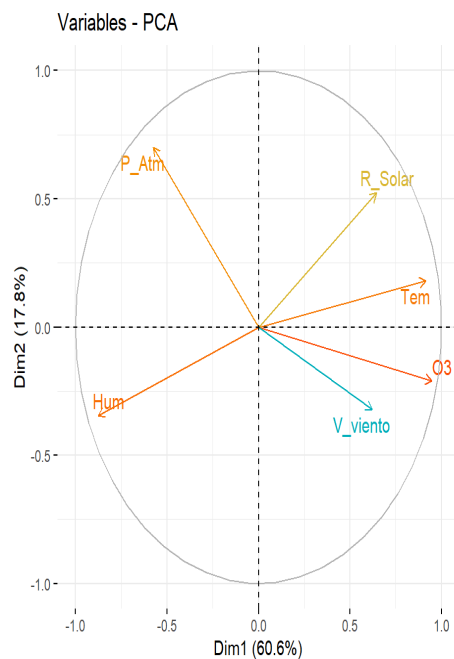
### Análisis de Componentes Principales (ACP)

El análisis de componentes principales conllevó a la reducción de seis dimensiones a solo dos, debido a que estas dos dimensiones acumulan el 78.38 % de la varianza y valores propios mayores a uno. Siendo el ozono, la temperatura y la humedad quienes más aporta en la variación en la dimensión uno. Mientras que la dimensión uno cuenta con la presión atmosférica como mayor aportante.

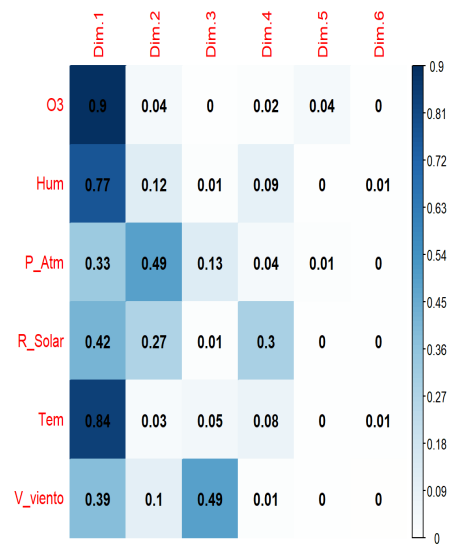
Variables	Dimensión					
	Dim. 1	Dim. 2	Dim. 3	Dim. 4	Dim. 5	Dim. 6
Ozono	0.898487	0.044529	2.6529e-06	0.017617	0.0393114	5.2351e-05
Humedad	0.769220	0.118419	1.4442e-02	0.091894	0.0001343	5.8891e-03
Presión Atmosférica	0.331244	0.492359	1.2643e-01	0.040361	0.0092312	3.6995e-04
Radiación Solar	0.415731	0.273767	1.1959e-02	0.295067	0.0034733	2.1853e-06
Temperatura	0.836419	0.032515	4.6892e-02	0.075493	0.0027858	5.8948e-03
Velocidad del Viento	0.386345	0.103762	4.9212e-01	0.014376	0.0033931	3.1027e-07
Valores Propios por Dimensión						
Valor Propio	3.637447	1.065353	0.691853	0.534809	0.058329	0.012208
Varianza (%)	60,6242	17,7558	11,5309	8,9135	0,9722	0,2035
Varianza Acumulada(%)	60,6242	<b>78,3800</b>	89,9109	98,8244	99,7965	100

Tabla **5-4**: Contribución de las variables a las dimensiones o componentes principales

En la figura **5-5a** confirma las correlaciones presentadas en la matriz de correlación (figura **5-3**) donde la temperatura, velocidad del viento y la radiación solar presentaron correlaciones positivas, mientras que la presión atmosférica y la humedad, correlaciones negativas. Estas direcciones se mantienen en la figura. Presentando un agrupamiento importante entre la temperatura, la velocidad del viento y el ozono. De acuerdo a esto, es posible que la estación esté dependiendo de la corriente de aire para tomar datos de los sectores aledaños que son los generadores de ozono. Aunque desconozco el lugar donde está la estación ambiental, es posible deducir según la figura, que se encuentra en un lugar donde está expuesto a calor constante, libre de sombra.



(a) Circulo de correlación



(b) Calidad en la representaciones de las variables

Figura 5-5: Análisis multivariado para los dos primeros componentes principales

La radiación juega un papel importante en la producción del ozono. Por lo tanto, la calidad de las observaciones puede ser afectadas por la falta de mantenimiento y este reduzca la capacidad de percibir la radiación solar. Este supuesto se manifiesta debido a que los días despejados están asociados con temperaturas altas y elevados niveles de radiación. A menos que se presente una fuente alterna de calor en los alrededores como condensadores de aires acondicionados o maquinaria.

La figura 5-5b representa la calidad con la que fue representados las variables en cada uno de los componentes principales. En la dimensión uno el ozono fue muy bien representado. Seguido de la temperatura y la humedad. Mientras que la dimensión dos la mejor representada fue la presión atmosférica. A pesar de tener una baja calidad entre la representación de la presión atmosférica y la radiación solar suman más de 0.6.

## Cartas de Control Estadístico

Se realizaron estudio univariado a las observaciones de ozono y multivariado a las seis variables seleccionadas a la base de datos, buscando posibles tomas de datos con alteraciones de carácter natural o preventivo.

## Cartas de Control univariada

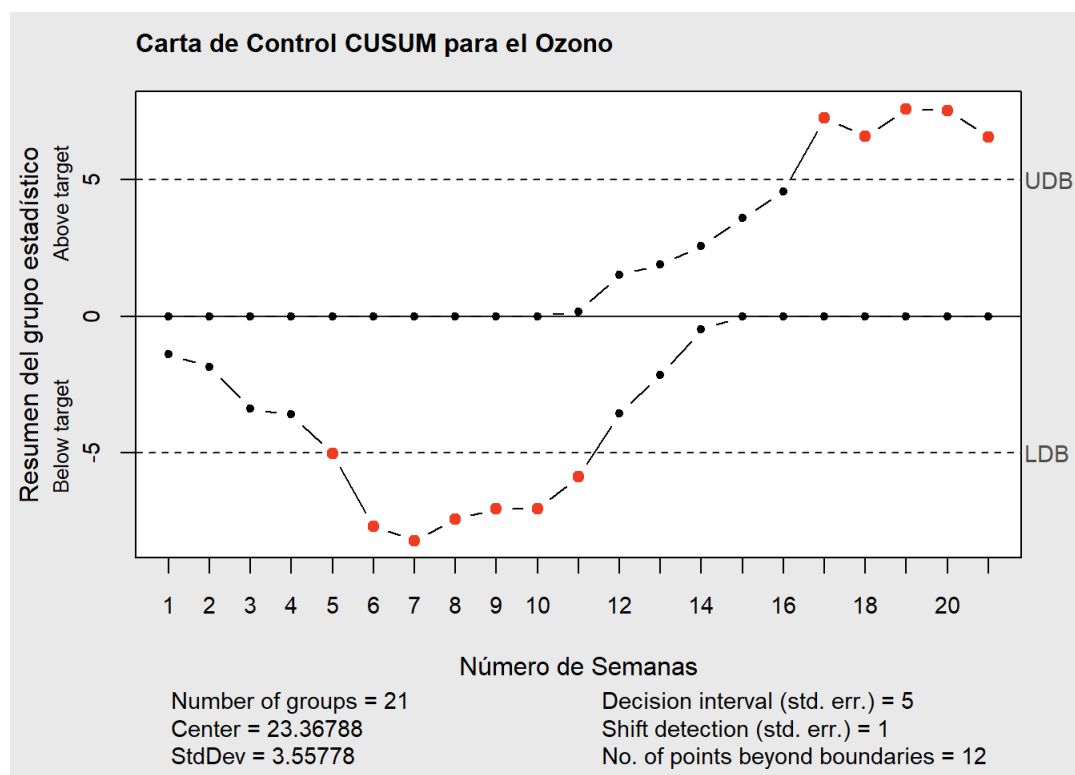


Figura 5-6: Carta de control estadístico CUSUM para el contaminante atmosférico Ozono

Buscando pequeños cambios en el registro de los datos ozono, se aplicó cartas de control CUSUM y EWMA. La figura 5-6 presenta un valor de desviación estándar en la figura de cinco y una permitancia de dos. Esta figura refleja como los datos presenta un descenso vertiginoso en las observaciones de ozono desde la semana tres hasta la siete. Con un posterior ascenso hasta llegar a la semana 21. Sin embargo, solo se evidencia una salida de los límites de control desde la semana cinco a la 11 por debajo y de la 17 a la 21 por encima. Este comportamiento también se puede observar en la figura 5-7 donde la carta de control EWMA manifiesta valores fuera de control desde la semana tres hasta la semana nueve por debajo, y por encima desde la semana 17 a la 20. se emplea una memoria de 5 y un valor de desviación de 2.5. Estas gráficas dejan en claro como las observaciones se encuentran fuera de control.

La semana cinco está dominada por picos máximos de hasta 25 y 30  $\mu g/m^3$ . Presentado en los días uno y cinco, niveles por encima de los 70  $\mu g/m^3$  de ozono sobre el medio día, que de acuerdo a los registros son las horas con mayor presencia de ozono troposférico mientras que la semana 6 solo registró 48 observaciones correspondientes a dos días con un máximo alrededor de 35  $\mu g/m^3$ . Para la semana 7 solo se registraron valores correspondientes a tres días y 12 horas, se presentó dos días con niveles entre 60  $\mu g/m^3$  y 80  $\mu g/m^3$ . Existe

un dato con un comportamiento que no se ajusta a lo necesario para formar ozono y es el registro que sucedió el 12 de agosto de 2018 entre la 1:00 am y las 1:59 am, con una concentración de  $37\mu g/m^3$ . A pesar que se encuentra por debajo de los límites permitidos en la normatividad queda un gran interrogante, debido a que en esa hora no se cuenta con los factores que promueven la reacción fotoquímica para la producción del ozono. La semana ocho normaliza la toma de datos al igual que la nueve. Presentado niveles máximos entre  $60\mu g/m^3$  y  $70\mu g/m^3$ . A esto se le suma dos días con ascensos en los valores en horas cercanas a la media noche. La semana 9 presenta niveles máximos por día de aproximadamente  $87\mu g/m^3$  y en días con bajos niveles de  $43\mu g/m^3$ . Se presentó un repunte sobres las 8:00 pm del 26 de agosto, lo cual puede estar asociado al aumento de la velocidad del viento, evidenciando que se presentan sectores de mayor acumulación en los alrededores de la estación ambiental donde los registros pudiesen llegar a ser más altos de los registrados en la ubicación actual de la estación.

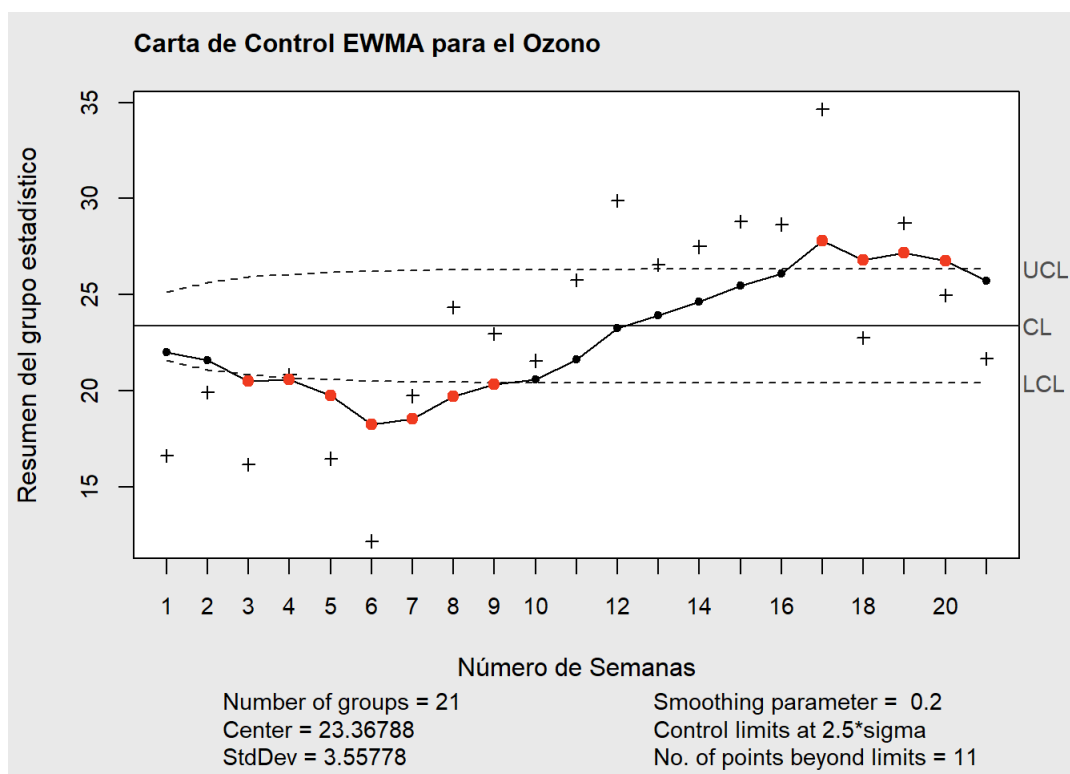


Figura 5-7: Carta de control estadístico EWMA para el contaminante atmosférico Ozono

En la semana 17 las observaciones de ozono se incrementan como se evidencia en la figura 5-6. Durante los tres días que registró datos, se presentó niveles máximos de hasta  $95.6\mu g/m^3$  como también niveles mínimos de  $5.9\mu g/m^3$ . La semana 23 (en la figura 5-6 y 5-7 como 18) inicia las observaciones después de un periodo acumulado cercano a las 6 semanas, manifestando un rango entre los valores máximos de los tres días de  $30\mu g/m^3$ .

Con un valor máximo aproximado de  $64 \mu\text{g}/\text{m}^3$ . La semana 24 (en la figura 5-6 y 5-7 como 19) indica un rango entre los valores máximos de los siete días de  $29.62 \mu\text{g}/\text{m}^3$ . Con un valor máximo aproximado de  $79.81 \mu\text{g}/\text{m}^3$ . Se observa dos días con inusuales incrementos en el valor de los datos en horas de la madrugada. La semana 25 (en la figura 5-6 y 5-7 como 20) indica un rango entre los valores máximos de los siete días de  $33.33 \mu\text{g}/\text{m}^3$ . Con un valor máximo aproximado de  $86.83 \mu\text{g}/\text{m}^3$ . Con la presencia de un día con inusuales incrementos en el valor de las observaciones en horas de la madrugada. La semana 26 (en la figura 5-6 y 5-7 como 21) indica un rango entre los valores máximos de los siete días de  $30.3 \mu\text{g}/\text{m}^3$ . Con un valor máximo aproximado de  $6.88 \mu\text{g}/\text{m}^3$ . Se observa dos días con inusuales incrementos en el valor de los datos en horas de la madrugada. Los inusuales incrementos en los valores del ozono que se encuentran entre el intervalo de las 11:00 pm y 5:00 am del día siguiente, pueden estar asociados a causas naturales o industriales. El ozono puede estar presente en las zonas urbanas debido a dos factores: El primero reacciones fotoquímicas a partir de metano, monóxido de nitrógeno y monóxido de carbono. El segundo por complejos mecanismos meteorológicos que transportan ozono estratosférico a la troposfera [Torres \(1991\)](#). Sin embargo, se puede asociar la presencia del ozono a labores industriales que buscan la desinfección de superficies, tratamientos de agua, alimentos, conductos de ventilación o lavanderías de tipo industrial.

## Cartas de Control Multivariado

### Carta de Control en Base a Análisis de Componentes Principales

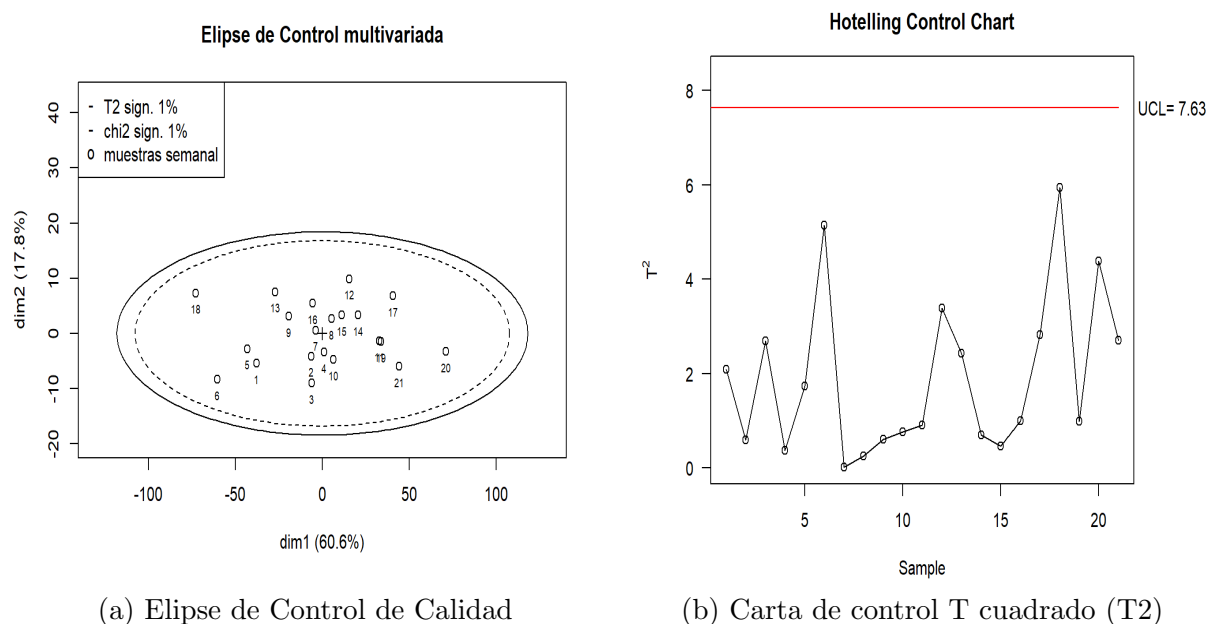


Figura 5-8: Control de calidad multivariado para los dos primeros componentes principales

De acuerdo a la figura **5-8a** las dimensiones uno y dos obtenidas de las cinco variables que tienen correlación con el ozono, no presenta observaciones exteriores a la elipse de control de calidad. Este resultado es corroborado por la figura **5-8b**, al evidenciar valores por debajo de la línea roja. Al realizar el control de calidad multivariado, se puede observar como las condiciones meteorológicas influyen directamente en la formación de ozono. Por lo tanto, se puede afirmar que las dos dimensiones obtenidas del análisis de componentes principales, se encuentran bajo control. De solo emplear el análisis univariado para el ozono y determinada la correlación que existe entre este y las variables meteorológicas, se cometería el error tipo I al señalar la variable ozono fuera de control y no tener en cuenta el efecto o acción que ejercen las demás variables.



## Capítulo 6

### Conclusiones

- Se logra realizar la depuración de la base de datos pasando de 4455 a 2807 datos al remover los registros en cero del ozono y los valores de las variables meteorológicas que lo acompañaban. Esta remoción de ceros, eliminó el registro de seis semanas. Luego se organiza por semanas y se realiza muestreo aleatorio de acuerdo a la cantidad de datos presentes. Como resultado se obtiene una base de datos con 48 observaciones por semana para un total de 21 semanas. El ozono como contaminante atmosférico y vigilado por la estación ambiental Catumare entre las 0:00 horas del 25 de junio hasta las 13:59 horas del 31 de diciembre de 2018, presenta claras inconsistencias en el registro de datos. Dificulta un análisis profundo de los datos para un periodo de tiempo superior a tres meses. Durante la depuración de datos se logra identificar las diferencias entre los diversos equipos empleados para la recolección de los mismos, donde algunos métodos (meteorológicos) registran datos constantes, en tanto, el ozono registra solo un 64 % observaciones diferentes de cero por periodos prolongados.
- Al aplicar análisis de componentes principales a la base de datos, se logró reducir de seis a dos variables que justificaran el 78 % de los datos. De esta manera se identificó correlaciones entre el ozono, la humedad, la temperatura, velocidad del viento, radiación solar y presión atmosférica. En el caso de la temperatura y la radiación solar se asocia con día soleados que es un ambiente propicio para la generación de ozono. Mientras que los días nublados se asocia con el aumento de la humedad y factores que favorecen el aumento de la humedad relativa como es el caso de la presión atmosférica. El viento favorece el desplazamiento de los contaminantes desde los lugares de producción a la estación ambiental.
- Al realizar un estudio univariado con cartas de control de calidad se identifica un incremento en la concentración de ozono en horas de la madrugada. A pesar que a inicio de la noche los sensores registran aumentos en la concentración de ozono



por el aumento que se genera en la velocidad del viento, los niveles no superan los registrados antes de las 18:30 horas. El incremento en la madrugada puede estar asociados a actividades industriales alrededor de la estación. También permite definir un proceso fuera de control al presentar múltiples valores por debajo y por encima de los límites calculados. Sin embargo, al realizar un análisis en el que se involucra todas las variables, se determina que el proceso no se encuentra fuera de control y que las semanas fuera de control del análisis univariado se deben a unas variaciones aleatorias o naturales, por lo que solo el uso del análisis univariado conllevaría a tomar un error en la decisión tipo I.

# Referencias

- Alberto, W. D., del Pilar, D. M., Valeria, A. M., Fabiana, P. S., Cecilia, H. A., y de los Ángeles, B. M. (2001). Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality. a case study:: Suquía river basin (córdoba argentina). *Water research*, 35(12), 2881–2894.
- Balcázar Gonzaga, D. V. (2013). *Análisis del nivel de contaminación atmosférica ocasionado por la emisión de gases de escape del parque automotor en la ciudad de Loja* (B.S. thesis). Universidad del Azuay.
- Borsese, A., y Esteban, S. (2005). Química, educación ambiental y vida cotidiana: el ozono troposférico. *Enseñanza de las Ciencias*, 23(2), 251–261.
- Canales, F. H. d., Alvarado, E. L. d., y Pineda, E. B. (1994). *Metodología de la investigación: Manual para el desarrollo de personal de salud* (n.º 35). Organización Panamericana de la Salud. PALTEX.
- de Ambiente y Desarrollo Sostenible, M. (2017, nov). *Resolución 2254 de 2017*.
- de Colombia, C. P. (1991). Constitución política de colombia. recuperado de: <http://www.registraduria.gov.co/IMG/pdf/constitucio-politica-colombia-1991.pdf>.
- de la Salud, O. M. (2016). *Impacto del medio ambiente en la salud*. World Health Organization.
- Fenger, J. (1999). Urban air quality. *Atmospheric environment*, 33(29), 4877–4900.
- Gramsch, E., Cereceda-Balic, F., Oyola, P., y Von Baer, D. (2006). Examination of pollution trends in santiago de chile with cluster analysis of pm10 and ozone data. *Atmospheric environment*, 40(28), 5464–5475.
- Hernández-Sampieri, R., Fernández-Collado, R., y Baptista-Lucio, P. (2017). Selección de la muestra.
- Ibarra-Berastegi, G., Elias, A., Barona, A., Saenz, J., Ezcurra, A., y de Argandoña, J. D. (2008). From diagnosis to prognosis for forecasting air pollution using neural networks: Air pollution monitoring in bilbao. *Environmental Modelling & Software*, 23(5), 622–637.
- Ibarra-Berastegi, G., Sáenz, J., Ezcurra, A., Ganzedo, U., de Argandoña, J. D., Errasti, I., ... Polanco-Martínez, J. (2009). Assessing spatial variability of so2 field as detected by an air quality network using self-organizing maps, cluster, and principal

- component analysis. *Atmospheric Environment*, 43(25), 3829–3836.
- IDEAM. (2018). *Informe del estado de la calidad del aire en colombia 2017*. Instituto de Hidrología, Meteorología y Estudios Ambientales Bogotá DC.
- López, P. L. (2004). Población muestra y muestreo. *Punto cero*, 9(08), 69–74.
- Lozares Colina, C., y López-Roldán, P. (1991). El análisis multivariado: definición, criterios y clasificación. *Papers: revista de sociología*(37), 009–29.
- Martínez, J. M. P. (2016). El papel del análisis por componentes principales en la evaluación de redes de control de la calidad del aire. *Comunicaciones en Estadística*, 9(2), 271–294.
- Mata, M. (1994). Cómo conocer la audiencia de una emisora, los sondeos de audiencia. *Cuadernos de investigación. Quito-Ecuador: Aler*.
- Medrano, S. (2003). Medición de humedad relativa con psicrómetro. *Boletín periódico del laboratorio de metrología. México*.
- Montgomery, D. C. (2012). *Statistical quality control*. Wiley Global Education.
- Moore, D. S. (2005). *Estadística aplicada básica*. Antoni Bosch editor.
- Morillas, A. (2007). Muestreo en poblaciones finitas. *Recuperado de: <http://webpersonal.uma.es/~morillas/muestreo.pdf> (07/11/2017)*.
- Nesselroade, J. R., y Cattell, R. B. (2013). *Handbook of multivariate experimental psychology*. Springer Science & Business Media.
- Organization, W. H. (2006). *Air quality guidelines: global update 2005: particulate matter, ozone, nitrogen dioxide, and sulfur dioxide*. Autor.
- Organization, W. H., y cols. (2006). *Who air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide: global update 2005: summary of risk assessment* (Inf. Téc.). Geneva: World Health Organization.
- Organization, W. H., y cols. (2016). *Ambient air pollution: A global assessment of exposure and burden of disease*.
- Phillips, D. (2000). L'ozono: Protettore o inquinante? *LA CHIMICA NELLA SCUOLA*.
- Pires, J., Pereira, M., Alvim-Ferraz, M., y Martins, F. (2009). Identification of redundant air quality measurements through the use of principal component analysis. *Atmospheric environment*, 43(25), 3837–3842.
- Pires, J., Sousa, S., Pereira, M., Alvim-Ferraz, M., y Martins, F. (2008). Management of air quality monitoring using principal component and cluster analysis—part i: So<sub>2</sub> and pm<sub>10</sub>. *Atmospheric Environment*, 42(6), 1249–1260.
- Ruiz Barzola, O. (2013). *Gráficos de control de calidad multivariantes con dimensión variable* (Tesis Doctoral no publicada).
- Shrestha, S., y Kazama, F. (2007). Assessment of surface water quality using multivariate statistical techniques: A case study of the fuji river basin, japan. *Environmental Modelling & Software*, 22(4), 464–475.
- Singh, K. P., Malik, A., Mohan, D., y Sinha, S. (2004). Multivariate statistical techniques

- for the evaluation of spatial and temporal variations in water quality of gomti river (india)—a case study. *Water research*, 38(18), 3980–3992.
- Torres, J. (1991). *Construcción y aplicación de trayectorias de retroceso de parcelas de aire al estudio de la contaminación atmosférica por ozono en la zona metropolitana de la ciudad de México* (Tesis Doctoral no publicada). Tesis Maestría, División de Estudios de Posgrado, Facultad de Ingeniería . . . .
- Vargas Marcos, F. (2005). *La contaminación ambiental como factor determinante de la salud*. SciELO Public Health.
- Vecchi, R., Chiari, M., D'Alessandro, A., Fermo, P., Lucarelli, F., Mazzei, F., . . . others (2008). A mass closure and pmf source apportionment study on the sub-micron sized aerosol fraction at urban sites in Italy. *Atmospheric Environment*, 42(9), 2240–2253.
- Velásquez Martínez, L. J., y cols. (2019). Control estadístico de calidad para el diseño de indicadores al proceso de ventas en una industria del sector automotriz.
- Young, B., Young, K. R., y Josse, C. (2011). Vulnerability of tropical andean ecosystems to climate change. *Climate Change and Biodiversity in the Tropical Andes. SCOPE, IAI*, 170–181.